tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56° C), and supplemented with about 10 µg/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin. The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the American Type Culture Collection, Rockville, Maryland. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands and colleagues (*Gastroenterology* **80:**225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the follistatin-3 antigen.

Alternatively, additional antibodies capable of binding to the follistatin-3 antigen may be produced in a two-step procedure through the use of anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and that, therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, follistatin-3 -specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the follistatin-3 -specific antibody can be blocked by the follistatin-3 antigen. Such antibodies comprise anti-idiotypic antibodies to the follistatin-3 -specific antibody and can be used to immunize an animal to induce formation of further follistatin-3 -specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, follistatin-3 -binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For *in vivo* use of anti-Follistatin-3 in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art (Morrison, *Science* **229:**1202 (1985); Oi, *et al., BioTechniques* **4:**214 (1986); Cabilly, *et al.*, U.S. Patent No. 4,816,567; Taniguchi, *et al.*, EP 171496; Morrison, *et al.*, EP 173494; Neuberger, *et al.*, WO 8601533; Robinson, *et al.*, WO 8702671; Boulianne, *et al., Nature* **312:**643 (1984); Neuberger, *et al., Nature* **314:**268 (1985).

Reproductive System- and Cell Growth and Differentiation-Related Disorders

Diagnosis

The present inventors have discovered that follistatin-3 is expressed not only in Hodgkin's Lymphoma, but also in synovial fibroblasts, gall bladder, resting and serum-induced smooth muscle, testes, Merkel cells, HEL cells, hippocampus, TNF-α- and IFN-induced epithelial cells, keratinocyte, amygdala depression, HL-60 cells, hepatoma, progesterone-treated epidermal cells, endothelial cells, HSC172 cells, epithelioid sarcoma, activated T-cells, breast lymph node, pancreatic carcinoma, fetal dura mater, fetal lung, epididymis, placenta, dendritic cells, rejected kidney, and uterine cancer. For a number of reproductive system-related disorders and disorders related to the regulation of cell growth and differentiation, substantially altered (increased or decreased) levels of follistatin-3 gene expression can be detected in reproductive system tissue or other cells or bodily fluids (e.g., sera, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to a "standard" follistatin-3 gene expression level, that is, the follistatin-3 expression level in reproductive system tissues or bodily fluids from an individual not having the reproductive system or cell growth and differentiation disorder. Thus, the invention provides a diagnostic method useful during diagnosis of a reproductive system or cell growth and differentiation disorder, which involves measuring the expression level of the gene encoding the follistatin-3 polypeptide in reproductive system tissue or other cells or body fluid from an individual and comparing the measured gene expression level with a standard follistatin-3 gene expression level, whereby an increase or decrease in the gene expression level compared to the standard is indicative of a reproductive or cell growth and differentiation system disorder.

In particular, it is believed that certain tissues in mammals with cancer of various cells and tissues of the reproductive or other systems express significantly reduced levels of the follistatin-3 polypeptide and mRNA encoding the follistatin-3 polypeptide when compared to a corresponding "standard" level. Further, it is believed that enhanced levels of the follistatin-3 polypeptide can be detected in certain body fluids (e.g., sera, plasma, urine, and spinal fluid) from mammals with such a cancer when compared to sera from mammals of the same species not having the cancer.

Thus, the invention provides a diagnostic method useful during diagnosis of reproductive system or cell growth and differentiation disorders, including cancers of these systems, which involves measuring the expression level of the gene encoding the follistatin-3 polypeptide in reproductive system tissue or other cells or body fluid from

an individual and comparing the measured gene expression level with a standard follistatin-3 gene expression level, whereby an increase or decrease in the gene expression level compared to the standard is indicative of a reproductive system disorder or a disorder of the regulation of cell growth and differentiation.

Where a diagnosis of a disorder in the reproductive or other system including diagnosis of a tumor, has already been made according to conventional methods, the present invention is useful as a prognostic indicator, whereby patients exhibiting depressed follistatin-3 gene expression will experience a worse clinical outcome relative to patients expressing the gene at a level nearer the standard level.

By "assaying the expression level of the gene encoding the follistatin-3 polypeptide" is intended qualitatively or quantitatively measuring or estimating the level of the follistatin-3 polypeptide or the level of the mRNA encoding the follistatin-3 polypeptide in a first biological sample either directly (e.g., by determining or estimating absolute polypeptide level or mRNA level) or relatively (e.g., by comparing to the follistatin-3 polypeptide level or mRNA level in a second biological sample). Preferably, the follistatin-3 polypeptide level or mRNA level in the first biological sample is measured or estimated and compared to a standard follistatin-3 polypeptide level or mRNA level, the standard being taken from a second biological sample obtained from an individual not having the disorder or being determined by averaging levels from a population of individuals not having a disorder of the reproductive system or of regulation of cell growth and differentiation. As will be appreciated in the art, once a standard follistatin-3 polypeptide level or mRNA level is known, it can be used repeatedly as a standard for comparison.

By "biological sample" is intended any biological sample obtained from an individual, body fluid, cell line, tissue culture, or other source which contains follistatin-3 polypeptide or mRNA. As indicated, biological samples include body fluids (such as sera, plasma, urine, synovial fluid and spinal fluid) which contain free follistatin-3 polypeptide, reproductive system tissue, and other tissue sources found to express complete or mature follistatin-3 or a follistatin-3 receptor. Methods for obtaining tissue biopsies and body fluids from mammals are well known in the art. Where the biological sample is to include mRNA, a tissue biopsy is the preferred source.

The present invention is useful for diagnosis or treatment of various reproductive system-related disorders and disorders of the regulation of cell growth and differentiation in mammals, preferably humans. Such disorders include tumors, cancers, interstitial lung disease, and any disregulation of the growth and differentiation patterns of cell function including, but not limited to, autoimmunity, arthritis,

leukemias, lymphomas, immunosuppression, immunity, humoral immunity, inflammatory bowel disease, myelosuppression and the like.

Total cellular RNA can be isolated from a biological sample using any suitable technique such as the single-step guanidinium-thiocyanate-phenol-chloroform method described by Chomczynski and Sacchi (*Anal. Biochem.* **162**:156-159 (1987)). Levels of mRNA encoding the follistatin-3 polypeptide are then assayed using any appropriate method. These include Northern blot analysis, S1 nuclease mapping, the polymerase chain reaction (PCR), reverse transcription in combination with the polymerase chain reaction (RT-PCR), and reverse transcription in combination with the ligase chain reaction (RT-LCR).

Assaying follistatin-3 polypeptide levels in a biological sample can occur using antibody-based techniques. For example, follistatin-3 polypeptide expression in tissues can be studied with classical immunohistological methods (Jalkanen, M., *et al.*, *J. Cell. Biol.* **101**:976-985 (1985); Jalkanen, M., *et al.*, *J. Cell. Biol.* **105**:3087-3096 (1987)). Other antibody-based methods useful for detecting follistatin-3 polypeptide gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (^{125}I, ^{121}I), carbon (^{14}C), sulfur (^{35}S), tritium (^{3}H), indium (^{112}In), and technetium (^{99m}Tc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying follistatin-3 polypeptide levels in a biological sample obtained from an individual, follistatin-3 polypeptide can also be detected *in vivo* by imaging. Antibody labels or markers for *in vivo* imaging of follistatin-3 polypeptide include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A follistatin-3 polypeptide-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, ^{131}I, ^{112}In, ^{99m}Tc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously or intraperitoneally) into the mammal to be examined for immune system disorder. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected

will normally range from about 5 to 20 millicuries of ^{99m}Tc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain follistatin-3 polypeptide. *In vivo* tumor imaging is described by Burchiel and coworkers (Chapter 13 in *Tumor Imaging: The Radiochemical Detection of Cancer*, Burchiel, S. W. and Rhodes, B. A., eds., Masson Publishing Inc. (1982)).

Treatment

As noted above, follistatin-3 polynucleotides and polypeptides are useful for diagnosis of conditions involving abnormally high or low expression of follistatin-3 activities. Given the cells and tissues where follistatin-3 is expressed as well as the activities modulated by follistatin-3, it is readily apparent that a substantially altered (increased or decreased) level of expression of follistatin-3 in an individual compared to the standard or "normal" level produces pathological conditions related to the bodily system(s) in which follistatin-3 is expressed and/or is active.

It will also be appreciated by one of ordinary skill that, since the follistatin-3 polypeptide of the invention is a member of the inhibin-related protein family the mature secreted form of the protein may be released in soluble form from the cells which express the follistatin-3 by proteolytic cleavage. Therefore, when follistatin-3 mature form is added from an exogenous source to cells, tissues or the body of an individual, the protein will exert its physiological activities on its target cells of that individual.

Therefore, it will be appreciated that conditions caused by a decrease in the standard or normal level of Follistatin-3 activity in an individual, particularly disorders of the reproductive system, can be treated by administration of follistatin-3 polypeptide (in the form of the mature protein). Thus, the invention also provides a method of treatment of an individual in need of an increased level of follistatin-3 activity comprising administering to such an individual a pharmaceutical composition comprising an amount of an isolated follistatin-3 polypeptide of the invention, particularly a mature form of the follistatin-3 protein of the invention, effective to increase the follistatin-3 activity level in such an individual.

Follistatin-3 may be used to treat male sterility by its innate ability to bind activin and, as a result, prevent activin-binding to its receptor. Activin receptor-binding results in a suppression of FSH secretion. Increased levels of FSH, in turn, result in an increase in spermatogenesis (Ying, S.-Y. *Endocrine Rev.* **9**:267-293 (1988)). Thus, a decrease in the effective concentration of activin will result in an FSH-mediated increase in spermatogenesis. In addition, since activin elicits a number of biological effects including the modulation of gonadal androgen biosynthesis (Hsueh, A. J. W., *et al., Proc. Natl. Acad. Sci. USA* **84**:5082-5086 (1987)), the attenuation of growth hormone

secretion (Bilezikjian, L. M., *et al.*, *Endocrinology* **126:**2369-2376 (1990)), the promotion of erythroid cell differentiation (Eto, Y., *et al.*, *Biochem. Biophys. Res. Comm.* **142:**1095-1103 (1987)), the induction of mesoderm formation (Smith, J. C., *et al.*, *Nature* **345:**729-731 (1990)), and the maintenance of nerve cell survival (Schubert, D., *et al.*, *Nature* **344:**868-870 (1990)), and since follistatin-3 directly inhibits activin acitivity, follistatin-3 may be used to therapeutically regulate, as well as diagnostically evaluate, the conditions and events listed above. Follistatin-3 may also be used to inhibit the activin-induced differentiation of follicular granulosa cells (Nakamura, T., *et al.*, *Biochim. Biophys. Acta* **1135:**103-109 (1992)). Follistatin-3 may be used therapeutically to regulate autocrine endothelial cell activity and, as a result, induce angiogenesis (Kozian, D. H., *et al.*, *Lab. Invest.* **76:**267-276 (1997)). Follistatin-3 may also be used to inhibit the activity of activin and thereby prevent the observed activin-mediated inhibition of basal and androgen-stimulated proliferation and induction of apoptosis (Wang, Q. F., *et al.*, *Endocrinology* **137:**5476-5483 (1996)). Treatment to increase the expression or the presence of follistatin-3 may be used to inhibit the progression of gonadotroph adenomas, osteosarcomas, hepatomas, and other tumors and cancers including bone, breast, colon, lymphomas, leukemias, epithelial carcinomas, pancreatic, stomach, liver, lung, melanoma, prostate, ovarian, uterine, bladder, gliomas, retinoblastomas, sarcomas, and the like (Penabad, J. L., *et al.*, *J. Clin. Endocrinol. Metab.* **81:**3397-3403 (1996); Kato, M. V., *et al.*, *Oncogene* **12:**1361-1364 (1996)). Follistatin-3 may also be employed to stimulate wound healing. In this same manner, follistatin-3 may also be employed to treat other fibrotic disorders, including liver cirrhosis, osteoarthritis and pulmonary fibrosis. Follistatin-3 also increases the presence of eosinophils which have the distinctive function of killing the larvae of parasites that invade tissues, as in schistosomiasis, trichinosis and ascariasis. It may also be employed to regulate hematopoiesis, by regulating the activation and differentiation of various hematopoietic progenitor cells, for example, to release mature leukocytes from the bone marrow following chemotherapy, i.e., in stem cell mobilization. Follistatin-3 may also be employed to treat sepsis. Follistatin-3 may also be used to treat a number of disease states known to those of skill in the art which may be therapeutically regulated by exploiting the prohibitive interation of follistatin-3 with the activin molecule.

Formulations and Administration

The follistatin-3 polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with follistatin-3

polypeptide alone), the site of delivery of the follistatin-3 polypeptide composition, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" of follistatin-3 polypeptide for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of follistatin-3 polypeptide administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the follistatin-3 polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the follistatin-3 of the invention may be administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally, topically (as by powders, ointments, drops or transdermal patch), bucally, or as an oral or nasal spray. By "pharmaceutically acceptable carrier" is meant a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The follistatin-3 polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U., *et al.*, *Biopolymers* **22**:547-556 (1983)), poly (2- hydroxyethyl methacrylate; Langer, R., *et al., J. Biomed. Mater. Res.* **15**:167-277 (1981), and Langer, R., *Chem. Tech.* **12**:98-105 (1982)), ethylene vinyl acetate (Langer, R., *et al., Id.*) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release follistatin-3 polypeptide compositions also include liposomally entrapped follistatin-3 polypeptide. Liposomes containing follistatin-3 polypeptide are prepared by methods known in the art (DE 3,218,121; Epstein, *et al., Proc. Natl. Acad. Sci. (USA)* **82**:3688-3692 (1985); Hwang, *et al., Proc. Natl. Acad. Sci. (USA)* **77**:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S.

Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324). Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal follistatin-3 polypeptide therapy.

For parenteral administration, in one embodiment, the follistatin-3 polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the follistatin-3 polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The follistatin-3 polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of follistatin-3 polypeptide salts.

Follistatin-3 polypeptide to be used for therapeutic administration must be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic follistatin-3 polypeptide

compositions generally are placed into a container having a sterile access port. for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Follistatin-3 polypeptide ordinarily will be stored in unit or multi-dose containers. for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous follistatin-3 polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized follistatin-3 polypeptide using bacteriostatic water-for-injection (WFI).

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Agonists and Antagonists - Assays and Molecules

The invention also provides a method of screening compounds to identify those which enhance or block the action of follistatin-3 on cells, such as its interaction with follistatin-3-binding molecules such as activin, an activin-like molecule. or a follistatin-3 receptor molecule. An agonist is a compound which increases the natural biological functions of follistatin-3 or which functions in a manner similar to follistatin-3, while antagonists decrease or eliminate such functions.

In another aspect of this embodiment the invention provides a method for identifying an activin-like molecule or a receptor protein or other ligand-binding protein which binds specifically to a follistatin-3 polypeptide. For example. a cellular compartment, such as a membrane or a preparation thereof, may be prepared from a cell that expresses a molecule that binds follistatin-3. The preparation is incubated with labeled follistatin-3 and complexes of follistatin-3 bound to the activin-like molecule, receptor or other binding protein are isolated and characterized according to routine methods known in the art. Alternatively, the follistatin-3 polypeptide may be bound to a solid support so that binding molecules solubilized from cells are bound to the column and then eluted and characterized according to routine methods.

In the assay of the invention for agonists or antagonists, a cellular compartment, such as a membrane or a preparation thereof, may be prepared from a cell that expresses

a molecule that binds follistatin-3, such as a molecule of a signaling or regulatory pathway modulated by follistatin-3. The preparation is incubated with labeled follistatin-3 in the absence or the presence of a candidate molecule which may be a follistatin-3 agonist or antagonist. The ability of the candidate molecule to bind the binding molecule is reflected in decreased binding of the labeled ligand. Molecules which bind gratuitously, i.e., without inducing the effects of follistatin-3 on binding the follistatin-3 binding molecule, are most likely to be good antagonists. Molecules that bind well and elicit effects that are the same as or closely related to follistatin-3 are agonists.

Follistatin-3-like effects of potential agonists and antagonists may by measured, for instance, by determining activity of a second messenger system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect with that of follistatin-3 or molecules that elicit the same effects as follistatin-3. Second messenger systems that may be useful in this regard include but are not limited to AMP guanylate cyclase, ion channel or phosphoinositide hydrolysis second messenger systems.

Another example of an assay for follistatin-3 antagonists is a competitive assay that combines follistatin-3 and a potential antagonist with membrane-bound follistatin-3 receptor molecules or recombinant follistatin-3 receptor molecules under appropriate conditions for a competitive inhibition assay. Follistatin-3 can be labeled, such as by radioactivity, such that the number of follistatin-3 molecules bound to a receptor molecule can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a receptor molecule, without inducing follistatin-3-induced activities, thereby preventing the action of follistatin-3 by excluding follistatin-3 from binding.

Other potential antagonists include antisense molecules. Antisense technology can be used to control gene expression through antisense DNA or RNA or through triple-helix formation. Antisense techniques are discussed in a number of studies (for example, Okano, *J. Neurochem.* **56**:560 (1991); "Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression." CRC Press, Boca Raton, FL (1988)). Triple helix formation is discussed in a number of studies, as well (for instance, Lee, *et al., Nucleic Acids Research* **6**:3073 (1979); Cooney, *et al., Science* **241**:456 (1988); Dervan, *et*

al., Science **251**:1360 (1991)). The methods are based on binding of a polynucleotide to a complementary DNA or RNA. For example, the 5' coding portion of a polynucleotide that encodes the mature polypeptide of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of follistatin-3. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into follistatin-3 polypeptide. The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of follistatin-3.

The agonists and antagonists may be employed in a composition with a pharmaceutically acceptable carrier, e.g., as described above.

Antagonists of follistatin-3 may be employed, for instance, to treat a deficiency in FSH, estrogen, and other hormones. Follistatin-1 and follistatin-3 are potent inhibitors of FSH and estrogen production and secretion. As a result, a deficiency of these or related hormones may be corrected or ameliorated through the use of a follistatin-3 antagonist. A follistatin-3 antagonist may be used to prevent or inhibit or reduce the production of spermatozoa by inhibiting the interaction of follistatin-3 with activin. Antagonists of follistatin-3 may also be used to modulate gonadal androgen biosynthesis, attenuate growth hormone secretion, promote the differentiation of follicular granulosa, erythroid, and other cell types, induce mesoderm formation, and increase the survival of nerve cells. A follistatin-3 antagonist may be used to inhibit angiogenesis related to or independent of tumorigenesis. Follistatin-3 antagonists may also be useful in increasing the activity of activin and thereby increasing the observed activin-mediated inhibition of basal and androgen-stimulated proliferation and induction of apoptosis. Antagonists of follistatin-3 may be used to regulate the hormonal and growth factor environment, and consequently, the activity of macrophages and their precursors, and of neutrophils, basophils, B lymphocytes and some T-cell subsets, e.g., activated and CD8 cytotoxic T cells and natural killer cells, in certain auto-immune and chronic inflammatory and infective diseases. Examples of auto-immune diseases include multiple sclerosis, and insulin-dependent diabetes. The antagonists may also be employed to treat infectious diseases including silicosis, sarcoidosis, idiopathic pulmonary fibrosis by altering the activation state of mononuclear phagocytes. They may also be employed to treat idiopathic hyper-eosinophilic syndrome by preventing eosinophil production and activation. Endotoxic shock may also be treated by the antagonists by preventing the activation of macrophages. Any of the above antagonists

may be employed in a composition with a pharmaceutically acceptable carrier, e.g., as hereinafter described.

Gene Mapping

The nucleic acid molecules of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

In certain preferred embodiments in this regard, the cDNA herein disclosed is used to clone genomic DNA of a follistatin-3 gene. This can be accomplished using a variety of well known techniques and libraries, which generally are available commercially. The genomic DNA then is used for *in situ* chromosome mapping using well known techniques for this purpose.

In addition, in some cases, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis of the 3' untranslated region of the gene is used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Fluorescence *in situ* hybridization ("FISH") of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with probes from the cDNA as short as 50 or 60 bp (for a review of this technique, see Verma, *et al.*, *Human Chromosomes: A Manual Of Basic Techniques*, Pergamon Press, New York (1988)).

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, on the World Wide Web (McKusick, V. *Mendelian Inheritance In Man*, available on-line through Johns Hopkins University, Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in

some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Examples

Example 1(a): Expression and Purification of "His-tagged" Follistatin-3 in E. coli

The bacterial expression vector pHE-4 is used for bacterial expression in this example. pHE-4 encodes ampicillin antibiotic resistance ("Ampr") and contains a bacterial origin of replication ("ori"), an IPTG inducible promoter, a ribosome binding site ("RBS"), six codons encoding histidine residues that allow affinity purification using nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin sold by QIAGEN, Inc., *supra*, and suitable single restriction enzyme cleavage sites. These elements are arranged such that an inserted DNA fragment encoding a polypeptide expresses that polypeptide with the six His residues (i.e., a "6 X His tag") covalently linked to the amino terminus of that polypeptide.

The DNA sequence encoding the desired portion of follistatin-3 comprising the mature form of the follistatin-3 amino acid sequence is amplified from the deposited cDNA clone using PCR oligonucleotide primers which anneal to the amino terminal sequences of the desired portion of follistatin-3 and to sequences in the deposited construct 3' to the cDNA coding sequence. Additional nucleotides containing restriction sites to facilitate cloning in the pHE-4 vector are added to the 5' and 3' primer sequences, respectively.

For cloning the mature form of the follistatin-3 protein, the 5' primer has the sequence 5' TCA CGC CAT ATG GGC TCG GGG AAC C 3' (SEQ ID NO:12) containing the underlined *Nde* I restriction site followed by 16 nucleotides of the amino terminal coding sequence of the mature follistatin-3 sequence in SEQ ID NO:2. One of ordinary skill in the art would appreciate, of course, that the point in the protein coding sequence where the 5' primer begins may be varied to amplify a DNA segment encoding any desired portion of the complete follistatin-3 protein shorter or longer than the mature form of the protein. The 3' primer has the sequence 5' CAT CCG GGT ACC TTA TTA CAC GAA GTT CTC TTC CTC TTC TG 3' (SEQ ID NO:13) containing the underlined *Asp* 718 restriction site followed by two stop codons and 23

nucleotides complementary to the 3' end of the coding sequence of the follistatin-3 DNA sequence in Figures 1A, 1B, and 1C.

The amplified follistatin-3 DNA fragment and the vector pHE4 are digested with *Nde* I and *Asp* 718 and the digested DNAs are then ligated together. Insertion of the follistatin-3 DNA into the restricted pHE4 vector places the follistatin-3 protein coding region downstream from the IPTG-inducible promoter and in-frame with an initiating AUG and the six histidine codons.

The ligation mixture is transformed into competent *E. coli* cells using standard procedures such as those described by Sambrook and colleagues (*Molecular Cloning: a Laboratory Manual, 2nd Ed.*; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)). *E. coli* strain M15/rep4, containing multiple copies of the plasmid pREP4, which expresses the lac repressor and confers kanamycin resistance ("Kan^r"), is used in carrying out the illustrative example described herein. This strain, which is only one of many that are suitable for expressing follistatin-3 protein, is available commercially (QIAGEN, Inc., *supra*). Transformants are identified by their ability to grow on LB plates in the presence of ampicillin and kanamycin. Plasmid DNA is isolated from resistant colonies and the identity of the cloned DNA confirmed by restriction analysis, PCR and DNA sequencing.

Clones containing the desired constructs are grown overnight ("O/N") in liquid culture in LB media supplemented with both ampicillin (100 µg/ml) and kanamycin (25 µg/ml). The O/N culture is used to inoculate a large culture, at a dilution of approximately 1:25 to 1:250. The cells are grown to an optical density at 600 nm ("OD600") of between 0.4 and 0.6. Isopropyl-β-D-thiogalactopyranoside ("IPTG") is then added to a final concentration of 1 mM to induce transcription from the lac repressor sensitive promoter, by inactivating the lacI repressor. Cells subsequently are incubated further for 3 to 4 hours. Cells then are harvested by centrifugation.

The cells are then stirred for 3-4 hours at 4°C in 6M guanidine-HCl, pH 8. The cell debris is removed by centrifugation, and the supernatant containing the follistatin-3 is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist, 1995, QIAGEN, Inc., *supra*). Briefly the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes

of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the follistatin-3 is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins can be eluted by the addition of 250 mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80° C.

The following alternative method may be used to purify follistatin-3 expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells are harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells ware then lysed by passing the solution through a microfluidizer (Microfuidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 x *g* for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 x *g* centrifugation for 15 min., the pellet is discarded and the follistatin-3 polypeptide-containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 x *g*) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by

vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded follistatin-3 polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 μm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the follistatin-3 polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the follistatin-3 polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant follistatin-3 polypeptide exhibits greater than 95% purity after the above refolding and purification steps. No major contaminant bands are observed from Commassie blue stained 16% SDS-PAGE gel when 5 μg of purified protein is loaded. The purified protein is also tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

Example 2: Cloning and Expression of Follistatin-3 protein in a Baculovirus Expression System

In this illustrative example, the plasmid shuttle vector pA2 is used to insert the cloned DNA encoding complete protein, including its naturally associated secretory signal (leader) sequence, into a baculovirus to express the mature follistatin-3 protein, using standard methods as described by Summers and colleagues (*A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*, Texas

Agricultural Experimental Station Bulletin No. 1555 (1987)). This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as *Bam* HI, *Xba* I and *Asp* 718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors could be used in place of the vector above, such as pAc373, pVL941 and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, by Luckow and coworkers (*Virology* **170:**31-39 (1989)).

The cDNA sequence encoding the full length follistatin-3 protein in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence shown in SEQ ID NO:2, is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene. The 5' primer has the sequence 5' CAT CGC GGA TCC GCC ATC ATG CGT CCC GGG GCG CCA GGG C 3' (SEQ ID NO:14) containing the underlined *Bam* HI restriction enzyme site, an efficient signal for initiation of translation in eukaryotic cells (Kozak, M., *J. Mol. Biol.* **196:**947-950 (1987)), followed by 22 of nucleotides of the sequence of the complete follistatin-3 protein shown in Figure 1A, beginning with the AUG initiation codon. The 3' primer has the sequence 5' CAT CCG GGT ACC TCA CAC GAA GTT CTC TTC CTC TTC TG 3' (SEQ ID NO:15) containing the underlined *Asp* 718 restriction site followed by 23 nucleotides complementary to the 3' noncoding sequence in Figure 1A.

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with *Bam* HI and *Asp* 718 and again is purified on a 1% agarose gel. This fragment is designated herein F1.

The plasmid is digested with the restriction enzymes *Bam* HI and *Asp* 718 and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine

procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.) This vector DNA is designated herein "V1".

Fragment F1 and the dephosphorylated plasmid V1 are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Statagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria are identified that contain the plasmid with the human follistatin-3 gene by digesting DNA from individual colonies using *Bam* HI and *Asp* 718 and then analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing. This plasmid is designated herein pA2Follistatin-3.

Five µg of the plasmid pA2Follistatin-3 is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner and colleaguew (*Proc. Natl. Acad. Sci. USA* **84**:7413-7417 (1987)). One µg of BaculoGold™ virus DNA and 5 µg of the plasmid pA2Follistatin-3 are mixed in a sterile well of a microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27°C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27°C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith (*supra*). An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10). After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in

35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4°C. The recombinant virus is called V-Follistatin-3.

To verify the expression of the follistatin-3 gene Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus V-Follistatin-3 at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μCi of ^{35}S-methionine and 5 μCi ^{35}S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the mature form of the follistatin-3 protein, and thus the cleavage point and length of the naturally associated secretory signal peptide.

Follistatin-3 protein has been produced by the abovedescribed process in a baculovirus expression system. The resultant follistatin-3 polypeptide was isolated and C-terminal sequencing analysis was used to confirm the prediction that the N-terminal 26 amino acids of the full-length follistatin-3 polypeptide shown in Figures 1A, 1B, and 1C (and in SEQ ID NO:2) are cleaved and that the mature form of the follistatin-3 polypeptide begins with methionine-27 as the N-terminal residue according to the numbering scheme of Figures 1A, 1B, and 1C (which is identical to methionine-1 according to the numbering scheme of SEQ ID NO:2). Of course, it is important to remember that the observed mature form of a secreted protein may vary according to a number of factors as detailed above.

Example 3: Cloning and Expression of Follistatin-3 in Mammalian Cells

A typical mammalian expression vector contains the promoter element, which mediates the initiation of transcription of mRNA, the protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription can be achieved with the early and late promoters from SV40, the long

terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIV1 and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter). Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) and pBC12MI (ATCC 67109). Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the gene can be expressed in stable cell lines that contain the gene integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful to develop cell lines that carry several hundred or even several thousand copies of the gene of interest. Another useful selection marker is the enzyme glutamine synthase (GS; Murphy, *et al.*, *Biochem J.* **227**:277-279 (1991); Bebbington, *et al.*, *Bio/Technology* **10**:169-175 (1992)). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

The expression vectors pC1 and pC4 contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen, *et al.*, *Mol. Cel. Biol.* **5**:438-447 (1985)) plus a fragment of the CMV-enhancer (Boshart, *et al.*, *Cell* **41**:521-530 (1985)). Multiple cloning sites, e.g., with the restriction enzyme cleavage sites *Bam* HI, *Xba* I and *Asp* 718, facilitate the cloning of the gene of interest. The vectors contain in addition the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene.

Example 3(a): Cloning and Expression in COS Cells

The expression plasmid, pFollistatin-3HA, is made by cloning a portion of the cDNA encoding the mature form of the follistatin-3 protein into the expression vector pcDNAI/Amp or pcDNAIII (which can be obtained from Invitrogen, Inc.).

The expression vector pcDNAI/amp contains: (1) an *E. coli* origin of replication effective for propagation in *E. coli* and other prokaryotic cells; (2) an ampicillin resistance gene for selection of plasmid-containing prokaryotic cells; (3) an SV40 origin

of replication for propagation in eukaryotic cells; (4) a CMV promoter, a polylinker, an SV40 intron; (5) several codons encoding a hemagglutinin fragment (i.e., an "HA" tag to facilitate purification) followed by a termination codon and polyadenylation signal arranged so that a cDNA can be conveniently placed under expression control of the CMV promoter and operably linked to the SV40 intron and the polyadenylation signal by means of restriction sites in the polylinker. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein described by Wilson and colleagues (*Cell* **37**:767 (1984)). The fusion of the HA tag to the target protein allows easy detection and recovery of the recombinant protein with an antibody that recognizes the HA epitope. pcDNAIII contains, in addition, the selectable neomycin marker.

A DNA fragment encoding the complete follistatin-3 polypeptide is cloned into the polylinker region of the vector so that recombinant protein expression is directed by the CMV promoter. The plasmid construction strategy is as follows. The follistatin-3 cDNA of the deposited clone is amplified using primers that contain convenient restriction sites, such as described above for construction of vectors for expression of follistatin-3 in *E. coli*. Suitable primers include the following, which are used in this example. The 5' primer, containing the underlined *Bam* HI site, a Kozak sequence, an AUG start codon, and 22 nucleotides of the 5' coding region of the complete follistatin-3 polypeptide, has the following sequence: 5' CAT CGC GGA TCC GCC ACC ATG CGT CCC GGG GCG CCA GGG C 3' (SEQ ID NO:16). The 3' primer, containing the underlined *Asp* 718 and 23 of nucleotides complementary to the 3' coding sequence immediately before the stop codon, has the following sequence: 5' TCA CCG CTC GAG CAC GAA GTT CTC TTC CTC TTC TG 3' (SEQ ID NO:17).

The PCR amplified DNA fragment and the vector, pcDNAI/Amp, are digested with *Bam* HI and *Asp* 718 and then ligated. The ligation mixture is transformed into *E. coli* strain SURE (Stratagene Cloning Systems, La Jolla, CA 92037), and the transformed culture is plated on ampicillin media plates which then are incubated to allow growth of ampicillin resistant colonies. Plasmid DNA is isolated from resistant colonies and examined by restriction analysis or other means for the presence of the fragment encoding the complete follistatin-3 polypeptide

For expression of recombinant follistatin-3, COS cells are transfected with an expression vector, as described above, using DEAE-dextran, as described, for instance, by Sambrook and coworkers (*Molecular Cloning: a Laboratory Manual*, Cold Spring Laboratory Press, Cold Spring Harbor, New York (1989)). Cells are incubated under conditions for expression of follistatin-3 by the vector.

Expression of the follistatin-3-HA fusion protein is detected by radiolabeling and immunoprecipitation, using methods described in, for example Harlow and colleagues (*Antibodies: A Laboratory Manual, 2nd Ed.*; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988)). To this end, two days after transfection, the cells are labeled by incubation in media containing ^{35}S-cysteine for 8 hours. The cells and the media are collected, and the cells are washed and the lysed with detergent-containing RIPA buffer: 150 mM NaCl, 1% NP-40, 0.1% SDS, 1% NP-40, 0.5% DOC, 50 mM TRIS, pH 7.5, as described by Wilson and colleagues (*supra*). Proteins are precipitated from the cell lysate and from the culture media using an HA-specific monoclonal antibody. The precipitated proteins then are analyzed by SDS-PAGE and autoradiography. An expression product of the expected size is seen in the cell lysate, which is not seen in negative controls.

Example 3(b): Cloning and Expression in CHO Cells

The vector pC4 is used for the expression of follistatin-3 polypeptide. Plasmid pC4 is a derivative of the plasmid pSV2-dhfr (ATCC Accession No. 37146). The plasmid contains the mouse DHFR gene under control of the SV40 early promoter. Chinese hamster ovary- or other cells lacking dihydrofolate activity that are transfected with these plasmids can be selected by growing the cells in a selective medium (alpha minus MEM, Life Technologies) supplemented with the chemotherapeutic agent methotrexate. The amplification of the DHFR genes in cells resistant to methotrexate (MTX) has been well documented (see, e.g., Alt, F. W., *et al., J. Biol. Chem.* **253:**1357-1370 (1978); Hamlin, J. L. and Ma. C. *Biochem. et Biophys. Acta,* **1097:**107-143 (1990); Page, M. J. and Sydenham, M. A. *Biotechnology* **9:**64-68 (1991)). Cells grown in increasing concentrations of MTX develop resistance to the drug by overproducing the target enzyme, DHFR, as a result of amplification of the DHFR gene. If a second gene is linked to the DHFR gene, it is usually co-amplified and over-expressed. It is known in the art that this approach may be used to develop cell lines carrying more than 1,000 copies of the amplified gene(s). Subsequently, when the methotrexate is withdrawn, cell lines are obtained which contain the amplified gene integrated into one or more chromosome(s) of the host cell.

Plasmid pC4 contains for expressing the gene of interest the strong promoter of the long terminal repeat (LTR) of the Rouse Sarcoma Virus (Cullen, *et al., Mol. Cell. Biol.* **5:**438-447 (1985)) plus a fragment isolated from the enhancer of the immediate early gene of human cytomegalovirus (CMV; Boshart, *et al., Cell* **41:**521-530 (1985)).

Downstream of the promoter are the following single restriction enzyme cleavage sites that allow the integration of the genes: *Bam* HI, *Xba* I, and *Asp* 718. Behind these cloning sites the plasmid contains the 3' intron and polyadenylation site of the rat preproinsulin gene. Other high efficiency promoters can also be used for the expression, e.g., the human ß-actin promoter, the SV40 early or late promoters or the long terminal repeats from other retroviruses, e.g., HIV and HTLVI. Clontech's Tet-Off and Tet-On gene expression systems and similar systems can be used to express the follistatin-3 polypeptide in a regulated way in mammalian cells (Gossen, M., and Bujard, H. *Proc. Natl. Acad. Sci. USA* **89**:5547-5551 (1992)). For the polyadenylation of the mRNA other signals, e.g., from the human growth hormone or globin genes can be used as well. Stable cell lines carrying a gene of interest integrated into the chromosomes can also be selected upon co-transfection with a selectable marker such as gpt, G418 or hygromycin. It is advantageous to use more than one selectable marker in the beginning, e.g., G418 plus methotrexate.

The plasmid pC4 is digested with the restriction enzymes *Bam* HI and *Asp* 718 and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

The DNA sequence encoding the complete follistatin-3 polypeptide is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the desired portion of the gene. The 5' primer containing the underlined *Bam* HI site, a Kozak sequence, an AUG start codon, and 22 nucleotides of the 5' coding region of the complete follistatin-3 polypeptide, has the following sequence: 5' CAT CGC GGA TCC *GCC ACC* ATG CGT CCC GGG GCG CCA GGG C 3' (SEQ ID NO:18). The 3' primer, containing the underlined *Asp* 718 restriction site and 23 of nucleotides complementary to the 3' coding sequence immediately before the stop codon as shown in Figure 1A (SEQ ID NO:1), has the following sequence: 5' CAT CCG GGT ACC TCA CAC GAA GTT CTC TTC CTC TTC TG 3' (SEQ ID NO:19).

The amplified fragment is digested with the endonucleases *Bam* HI and *Asp* 718 and then purified again on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC4 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene are used for transfection. Five µg of the expression plasmid pC4 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner, *et al., supra*). The plasmid pSV2-neo

contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of metothrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μM, 2 μM, 5 μM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100-200 μM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Follistatin-3 protein has been produced by the abovedescribed process in a CHO cell expression system. The resultant follistatin-3 polypeptide was isolated and C-terminal sequencing analysis was used to confirm the prediction that the N-terminal 26 amino acids of the full-length follistatin-3 polypeptide shown in Figures 1A, 1B, and 1C (and in SEQ ID NO:2) are cleaved and that the mature form of the follistatin-3 polypeptide begins with methionine-27 as the N-terminal residue according to the numbering scheme of Figures 1A, 1B, and 1C (which is identical to methionine-1 according to the numbering scheme of SEQ ID NO:2). Of course, it is important to remember that the observed mature form of a secreted protein may vary according to a number of factors as detailed above.

Example 4: Tissue distribution of Follistatin-3 mRNA expression

Northern blot analysis was carried out to examine follistatin-3 gene expression in human tissues, using methods described by, among others, Sambrook and colleagues (*supra*). A cDNA probe containing the entire nucleotide sequence of the follistatin-3 protein (SEQ ID NO:1) was labeled with ^{32}P using the *redi*prime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe was purified using a CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe was then used to examine various human tissues for follistatin-3 mRNA.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) were obtained from Clontech and were examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots were mounted and exposed to film at -70°C overnight, and films developed according to standard procedures. The follistatin-3-specific probe recognized an mRNA species of approximately 2.6 kb in most tissues examined.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of all publications (including patents, patent applications, journal articles, laboratory manuals, books, or other documents) cited herein are hereby incorporated by reference.

Further, the Sequence Listing submitted herewith, and the Sequence Listing submitted with U. S. Provisional Application Serial No. 60/056,248, filed on August 29, 1997 (to which the present application claims benefit of the filing date under 35 U.S.C. § 119(e)), in both computer and paper forms are hereby incorporated by reference in their entireties.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 4, line 1

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet

Name of depositary institution
American Type Culture Collection

Address of depositary institution (*including postal code and country*)
10801 University Boulevard
Manassas, Virginia 20110-2209
United States of America

Date of deposit	Accession Number
August 8, 1997	209199

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications; e.g., "Accession Number of Deposit"*)

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer: Virginia L Lilly	Authorized officer

What Is Claimed Is:

1. An isolated nucleic acid molecule nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a nucleotide sequence encoding the follistatin-3 polypeptide having the complete amino acid sequence in SEQ ID NO:2 (i.e., positions -26 to 237 of SEQ ID NO:2);

(b) a nucleotide sequence encoding the follistatin-3 polypeptide having the complete amino acid sequence in SEQ ID NO:2 excepting the N-terminal methionine (i.e., positions -25 to 237 of SEQ ID NO:2);

(c) a nucleotide sequence encoding the predicted mature follistatin-3 polypeptide having the amino acid sequence at positions 1 to 237 in SEQ ID NO:2;

(d) a nucleotide sequence encoding the follistatin-3 polypeptide having the complete amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 209199;

(e) a nucleotide sequence encoding the follistatin-3 polypeptide having the complete amino acid sequence excepting the N-terminal methionine encoded by the cDNA clone contained in ATCC Deposit No. 209199;

(f) a nucleotide sequence encoding the mature follistatin-3 polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 209199; and

(g) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e) or (f) above.

2. The nucleic acid molecule of claim 1 wherein said polynucleotide has the complete nucleotide sequence in Figures 1A, 1B, and 1C (SEQ ID NO:1).

3. The nucleic acid molecule of claim 1 wherein said polynucleotide has the nucleotide sequence in Figures 1A, 1B, and 1C (SEQ ID NO:1) encoding the follistatin-3 polypeptide having the amino acid sequence in positions -26 to 237 of SEQ ID NO:2.

4. The nucleic acid molecule of claim 1 wherein said polynucleotide has the nucleotide sequence in Figures 1A, 1B, and 1C (SEQ ID NO:1) encoding the follistatin-3 polypeptide having the amino acid sequence in positions -25 to 237 of SEQ ID NO:2.

5. The nucleic acid molecule of claim 1 wherein said polynucleotide has the nucleotide sequence in Figures 1A, 1B, and 1C (SEQ ID NO:1) encoding the mature follistatin-3 polypeptide having the amino acid sequence from about 1 to about 237 in SEQ ID NO:2.

6. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a polypeptide comprising the amino acid sequence of residues n-237 of SEQ ID NO:2, where n is an integer in the range of -26-12;

(b) a nucleotide sequence encoding a polypeptide comprising the amino acid sequence of residues -26-m of SEQ ID NO:2, where m is an integer in the range of -26-m of 207 to 237;

(c) a nucleotide sequence encoding a polypeptide having the amino acid sequence consisting of residues n-m of SEQ ID NO:2, where n and m are integers as defined respectively in (a) and (b) above; and

(d) a nucleotide sequence encoding a polypeptide consisting of a portion of the complete follistatin-3 amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 209199 wherein said portion excludes from 1 to about 37 amino acids from the amino terminus of said complete amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 209199;

(e) a nucleotide sequence encoding a polypeptide consisting of a portion of the complete follistatin-3 amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 209199 wherein said portion excludes from 1 to about 20 amino acids from the carboxy terminus of said complete amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 209199; and

(f) a nucleotide sequence encoding a polypeptide consisting of a portion of the complete follistatin-3 amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 209199 wherein said portion include a combination of any of the amino terminal and carboxy terminal deletions in (d) and (e), above.

7. The nucleic acid molecule of claim 1 wherein said polynucleotide has the complete nucleotide sequence of the cDNA clone contained in ATCC Deposit No. 209199.

8. The nucleic acid molecule of claim 1 wherein said polynucleotide has the nucleotide sequence encoding the follistatin-3 polypeptide having the complete amino acid sequence excepting the N-terminal methionine encoded by the cDNA clone contained in ATCC Deposit No. 209199.

9. The nucleic acid molecule of claim 1 wherein said polynucleotide has the nucleotide sequence encoding the mature follistatin-3 polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 209199.

10. An isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide having a nucleotide sequence identical to a nucleotide sequence in (a), (b), (c), (d), (e), (f), or (g) of claim 1 wherein said polynucleotide which hybridizes does not hybridize under stringent hybridization conditions to a polynucleotide having a nucleotide sequence consisting of only A residues or of only T residues.

11. An isolated nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of an epitope-bearing portion of a follistatin-3 polypeptide having an amino acid sequence in (a), (b), (c), (d), (e), or (f) of claim 1.

12. The isolated nucleic acid molecule of claim 11, which encodes an epitope-bearing portion of a follistatin-3 polypeptide wherein the amino acid sequence of said portion is selected from the group of sequences in SEQ ID NO:2 consisting of: Leu-14 to Ala-20, Ser-46 to Ile-55, Gly-88 to Pro-97, Gly-113 to Leu-133, Arg-138 to Glu-146, Pro-177 to Thr-191, and Gly-219 to about Val-237.

13. A recombinant vector that contains the polynucleotide of claim 1.

14. A recombinant vector that contains the polynucleotide of claim 1 operably associated with a regulatory sequence that controls gene expression.

15. A genetically engineered host cell that contains the polynucleotide of claim 1.

16. A genetically engineered host cell that contains the polynucleotide of claim 1 operatively associated with a regulatory sequence that controls gene expression.

17. A method for producing a follistatin-3 polypeptide, comprising:

(a) culturing the genetically engineered host cell of claim 16 under conditions suitable to produce the polypeptide; and

(b) recovering said polypeptide.

18. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) the amino acid sequence of the full-length follistatin-3 polypeptide having the complete amino acid sequence shown in SEQ ID NO:2 (i.e., positions -26 to 237 of SEQ ID NO:2);

(b) the amino acid sequence of the full-length follistatin-3 polypeptide having the complete amino acid sequence shown in SEQ ID NO:2 excepting the N-terminal methionine (i.e., positions -25 to 237 of SEQ ID NO:2);

(c) the amino acid sequence of the predicted mature follistatin-3 polypeptide having the amino acid sequence at positions 1 to 237 in SEQ ID NO:2;

(d) the amino acid sequence of the full-length follistatin-3 polypeptide having the complete amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 209199;

(e) the amino acid sequence of the full-length follistatin-3 polypeptide having the complete amino acid sequence excepting the N-terminal methionine encoded by the cDNA clone contained in ATCC Deposit No. 209199; and

(f) the amino acid sequence of the mature follistatin-3 polypeptide having the amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 209199.

19. An isolated polypeptide comprising an epitope-bearing portion of the follistatin-3 protein, wherein said portion is selected from the group consisting of: a polypeptide comprising amino acid residues Leu-14 to Ala-20 of SEQ ID NO:2; a polypeptide comprising amino acid residues Ser-46 to Ile-55 of SEQ ID NO:2; a polypeptide comprising amino acid residues Gly-88 to Pro-97 of SEQ ID NO:2; a polypeptide comprising amino acid residues Gly-113 to Leu-133 of SEQ ID NO:2; a polypeptide comprising amino acid residues Arg-138 to Glu-146 of SEQ ID NO:2; a polypeptide comprising amino acid residues Pro-177 to Thr-191 of SEQ ID NO:2; and

a polypeptide comprising amino acid residues Gly-219 to Val-237 of SEQ ID NO:2.

20. An isolated antibody that binds specifically to a follistatin-3 polypeptide of claim 18.

21. An isolated nucleic acid molecule comprising a polynucleotide having a sequence at least 95% identical to a sequence selected from the group consisting of:

(a) the nucleotide sequence of SEQ ID NO:4;

(b) the nucleotide sequence of SEQ ID NO:5;

(c) the nucleotide sequence of SEQ ID NO:6;

(d) the nucleotide sequence of SEQ ID NO:7;

(e) the nucleotide sequence of SEQ ID NO:8;

(f) the nucleotide sequence of SEQ ID NO:9;

(g) the nucleotide sequence of SEQ ID NO:10;

(h) the nucleotide sequence of SEQ ID NO:11;

(i) the nucleotide sequence of a portion of the sequence shown in Figures 1A, 1B, and 1C (SEQ ID NO:1) wherein said portion comprises at least 50 contiguous nucleotides from nucleotide 1 to 500; and

(j) the nucleotide sequence of a portion of the sequence shown in Figures 1A, 1B, and 1C (SEQ ID NO:1) wherein said portion consists of nucleotides 100-500, 200-500, 300-500, 400-500, 100-400, 200-400, 300-400, 100-300, 200-300, 100-200, 100-2495, 250-2495, 500-2495, 1000-2495, 1500-2495, 2000-2495, 100-2000, 250-2000, 500-2000, 1000-2000, 1500-2000, 100-1500, 250-1500, 500-1500, 1000-1500, 100-1000, 250-1000, and 500-1000 of SEQ ID NO:1; and

(k) a nucleotide sequence complementary to any of the nucleotide sequences in (a) through (j) above.

22. The isolated nucleic acid of claim 1 which is fused to a polynucleotide encoding a heterologous polypeptide.

23. The isolated polypeptide of claim 18 which is fused to a heterologous polypeptide

24. A method for preventing, treating, or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 18.

25. A method for preventing, treating, or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of the nucleic acid of claim 1.

26. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject related to expression or activity of follistatin-3 comprising:

(a) determining the presence or absence of a mutation in the nucleic acid of claim 1;

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

27. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject related to expression or activity of follistatin comprising:

(a) determining the presence or amount of expression of the polypeptide of claim 18 in a biological sample;

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

28. A method of identifying compounds capable of enhancing or inhibiting a follistatin-3 activity comprising:

(a) contacting the polypeptide of claim 18, with a candidate compound; and

(b) assaying for activity.

Figure 1A
Follistatin-3

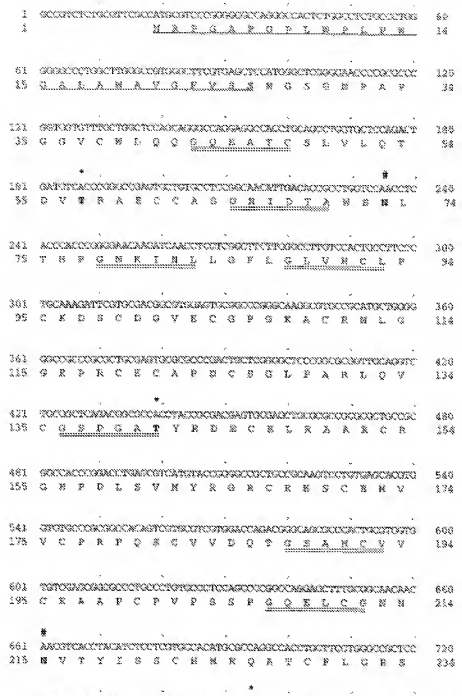
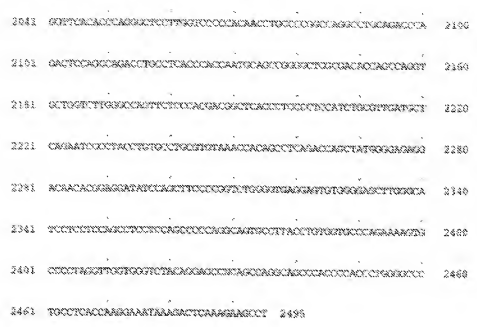


Figure 1B
Follistatin-3

Figure 1C
Follistatin-3



3 / 5

Figure 2

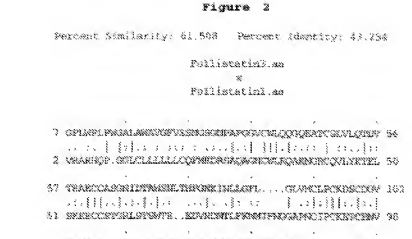
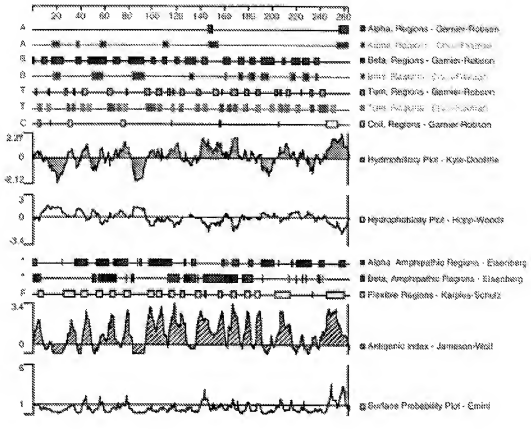

Percent Similarity: 61.508 Percent Identity: 43.254
Fpilistatin3.aa
x
Fpilistatin1.aa

Figure 3
Follistatin-3



SEQUENCE LISTING

<110> Duan, et al.

<120> Follistatin-3

<130> PF388

<140> Unassigned
<141> 1998-08-27

<150> 60/056,248
<151> 1997-08-29

<160> 19

<170> PatentIn Ver. 2.0

<210> 1
<211> 2495
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (19)..(807)

<220>
<221> mat_peptide
<222> (97)..(807)

<220>
<221> sig_peptide
<222> (19)..(96)

<400> 1
gccgtatctg cgttcgcc atg cgt ccc ggg gcg cca ggg cca ctc tgg ccc 51
Met Arg Pro Gly Ala Pro Gly Pro Leu Trp Pro
-25 -20

ctg ccc tgg ggg gcc ctg gct tgg gcc gtg ggc ttc gtg agc tcc atg 99
Leu Pro Trp Gly Ala Leu Ala Trp Ala Val Gly Phe Val Ser Ser Met
-15 -10 -5 -1 1

ggc tcg ggg aac ccc gcg ccc ggt ggc gtt tgc tgg ctc cag cag ggc 147
Gly Ser Gly Asn Pro Ala Pro Gly Gly Val Cys Trp Leu Gln Gln Gly
5 10 15

cag gag gcc acc tgc agc ctg gtg ctc cag act gat gtc acc cgg gcc 195
Gln Glu Ala Thr Cys Ser Leu Val Leu Gln Thr Asp Val Thr Arg Ala
20 25 30

gag tgc tgt gcc tcc ggc aac att gac acc gcc tgg tcc aac ctc acc 243
Glu Cys Cys Ala Ser Gly Asn Ile Asp Thr Ala Trp Ser Asn Leu Thr
35 40 45

cac ccg ggg aac aag atc aac ctc ctg ggc ttc ttg ggc ctt gtc cac 291
His Pro Gly Asn Lys Ile Asn Leu Leu Gly Phe Leu Gly Leu Val His
50 55 60 65

tgc ctt ccc tgc aaa gat tcg tgc gac ggc gtg gag tgc ggc ccg ggc 339
Cys Leu Pro Cys Lys Asp Ser Cys Asp Gly Val Glu Cys Gly Pro Gly
70 75 80

aag gcg tgc cgc atg ctg ggg ggc cgc ccg cgc tgc gag tgc gcg ccc 387
Lys Ala Cys Arg Met Leu Gly Gly Arg Pro Arg Cys Glu Cys Ala Pro

```
                95                  90                  95

gac tgc tcg ggg ctc ccg gcg cgg ttg cag gtc tgc ggc tca gac ggc   435
Asp Cys Ser Gly Leu Pro Ala Arg Leu Gln Val Cys Gly Ser Asp Gly
        100                 105                 110

gcc acc tac cgc gac gag tgc gag ctg cgc gcc gcg cgc tgc cgc ggc   483
Ala Thr Tyr Arg Asp Glu Cys Glu Leu Arg Ala Ala Arg Cys Arg Gly
    115                 120                 125

cac ccg gac ctg agc gtc atg tac cgg ggc cgc tgc cgc aag tcc tgt   531
His Pro Asp Leu Ser Val Met Tyr Arg Gly Arg Cys Arg Lys Ser Cys
130                 135                 140                 145

gag cac gtg gtg tgc ccg cgg cca cag tcg tgc gtc gtg gac cag acg   579
Glu His Val Val Cys Pro Arg Pro Gln Ser Cys Val Val Asp Gln Thr
                150                 155                 160

ggc agc gcc cac tgc gtg gtg tgt cga gcg gcg ccc tgc cct gtg ccc   627
Gly Ser Ala His Cys Val Val Cys Arg Ala Ala Pro Cys Pro Val Pro
            165                 170                 175

tcc agc ccc ggc cag gag ctt tgc ggc aac aac aac gtc acc tac atc   675
Ser Ser Pro Gly Gln Glu Leu Cys Gly Asn Asn Asn Val Thr Tyr Ile
        180                 185                 190

tcc tcg tgc cac atg cgc cag gcc acc tgc ttc ctg ggc cgc tcc atc   723
Ser Ser Cys His Met Arg Gln Ala Thr Cys Phe Leu Gly Arg Ser Ile
    195                 200                 205

ggc gtg cgc cac gcg ggc agc tgc gca ggc acc cct gag gag ccg cca   771
Gly Val Arg His Ala Gly Ser Cys Ala Gly Thr Pro Glu Glu Pro Pro
210                 215                 220                 225

ggt ggt gag tct gca gaa gag gaa gag aac ttc gag tgagcctgca         817
Gly Gly Glu Ser Ala Glu Glu Glu Glu Asn Phe Val
                230                 235

ggacaggcct ggccctggtg ccccaggccc cccatcatcc cctgttattt attgccacag 877

cagagtccaa tttatatgcm acggacactc cttagagccc ggattcggac cacttgggga 937

tcccagaccc tcccttgacga tatcctggaa ggactgagga agggaggcct gggggcmggc 997

tggtgggtgg gatagacctg cgttccggac actgagcgcc tgatttaggg ccctttccta 1057

ggatgcccca gcccctaccc taagacctat tgmcgggggag gattccacac ttccgctcct 1117

ttgggggata acctattaat tattgctatt atcaagaggg ctgggcattc tcrgctggta 1177

attcctgaag aggcatgact gcttttctca gccccaagcc tctagttctg gtgtgtacgg 1237

agggtctagc ctgggtgtgt acggagggtc tagcctgggt gagtacggag ggtctagcct 1297

gggtgaghac ggaggatcta gcctgggtga gtacggagag tctagcctgg gtgtgtatgg 1357

aggatctagc ctgggtgagt atggagggtc tagcctgggt gagtatggag ggtctagcct 1417

gggtgtgtac ggaggtcta gcctgggtga gtatggaggg tctagcctgg gtgtgtatgg 1477

agggtctagc ctgggtgagt atggagggtc tagcctgggt gtgtacggag ggtctagtct 1537

gagtgcgtgt ggggaacrca gaacactgtg accttagccc agcaagccag gccattcatg 1597

aaggccaaga aggctgccac catrcrctgc cagcccaaga actccagctt cccccactgm 1657

tctgcgtgcc ccttgcgtc ctgtgaagga cattgaycca tgcccagtgt gcccccrggg 1717
```

aaaggggcacg gcctgtgctc ctgaacaggg ctgtgcttgg ccacagaacc acccagcgtc 1777

tcccctgccg ctgttccagt cagttcatga ggcaacgtcg cgtggtttca gacgtggagc 1837

agccagcggc agctcagagc agggcactgt gtccggcgga gccaagtcca ctctgggggg 1897

gctctggccg ggaccacggg ccactgctca cccactggcc ccgagggggg tgtagcgccc 1957

aagactccag catgtgtgac atccggagtc ctggagccgg gtgtcccagt ggcaccacta 2017

ggtgccctgt gcctccacag tggggctcac acccagggct cctggtccc ccacaacctg 2077

cccctgccag gcctgcagac ccagactcca gccagacctg cctcaccccac caatgcagcc 2137

gtggctggcg acaccagccc ggtgctggtc ttgggccagt tctcccacga cgctcaccc 2197

tcccctccat ctgtgctgat gctcagatcc gcctacctgt gcctgctgt aaatacagc 2257

ctcagaccag ctatgtgggag aggacaccac ggaggatatc cagcttcccc ggtctgggct 2317

gaggaagtgtg ggtggcttgg gcatcctcct ccagcctcct ccagcccccg ggcagtgcct 2377

tacctgtggt gttcagaaaa gtgcccctag gttggtgggt ctacaggagc cncagccagg 2437

cagcccaccc cmrrctgggg ccctgcctca ccaaggaaat aaagactcaa agaaggcct 2495

<210> 2
<211> 263
<212> PRT
<213> Homo sapiens

<400> 2
Met Arg Pro Gly Ala Pro Gly Pro Leu Trp Pro Leu Pro Trp Gly Ala
-25 -20 -15

Leu Ala Trp Ala Val Gly Phe Val Ser Ser Met Gly Ser Gly Asn Pro
-10 -5 -1 1 5

Ala Pro Gly Gly Val Cys Trp Leu Gln Gln Gly Gln Glu Ala Thr Cys
10 15 20

Ser Leu Val Leu Gln Thr Asp Val Thr Arg Ala Glu Cys Cys Ala Ser
25 30 35

Gly Asn Ile Asp Thr Ala Trp Ser Asn Leu Thr His Pro Gly Asn Lys
40 45 50

Ile Asn Leu Leu Gly Phe Leu Gly Leu Val His Cys Leu Pro Cys Lys
55 60 65 70

Asp Ser Cys Asp Gly Val Glu Cys Gly Pro Gly Lys Ala Cys Arg Met
75 80 85

Leu Gly Gly Arg Pro Arg Cys Glu Cys Ala Pro Asp Cys Ser Gly Leu
90 95 100

Pro Ala Arg Leu Gln Val Cys Gly Ser Asp Gly Ala Thr Tyr Arg Asp
105 110 115

Glu Cys Glu Leu Arg Ala Ala Arg Cys Arg Gly His Pro Asp Leu Ser
120 125 130

Val Met Tyr Arg Gly Arg Cys Arg Lys Ser Cys Glu His Val Val Cys
135 140 145 150

Pro Arg Pro Gln Ser Cys Val Val Asp Gln Thr Gly Ser Ala His Cys

155 160 165

Val Val Cys Arg Ala Ala Pro Cys Pro Val Pro Ser Ser Pro Gly Gln
170 175 180

Glu Leu Cys Gly Asn Asn Asn Val Thr Tyr Ile Ser Ser Cys His Met
185 190 195

Arg Gln Ala Thr Cys Phe Leu Gly Arg Ser Ile Gly Val Arg His Ala
200 205 210

Gly Ser Cys Ala Gly Thr Pro Glu Glu Pro Pro Gly Gly Glu Ser Ala
215 220 225 230

Glu Glu Glu Glu Asn Phe Val
235

<210> 3
<211> 317
<212> PRT
<213> Homo sapiens

<400> 3
Met Val Arg Ala Arg His Gln Pro Gly Gly Leu Cys Leu Leu Leu Leu
1 5 10 15

Leu Leu Cys Gln Phe Met Glu Asp Arg Ser Ala Gln Ala Gly Asn Cys
20 25 30

Trp Leu Arg Gln Ala Lys Asn Gly Arg Cys Gln Val Leu Tyr Lys Thr
35 40 45

Glu Leu Ser Lys Glu Glu Cys Cys Ser Thr Gly Arg Leu Ser Thr Ser
50 55 60

Trp Thr Glu Glu Asp Val Asn Asp Asn Thr Leu Phe Lys Trp Met Ile
65 70 75 80

Phe Asn Gly Gly Ala Pro Asn Cys Ile Pro Cys Lys Glu Thr Cys Glu
85 90 95

Asn Val Asp Cys Gly Pro Gly Lys Lys Cys Arg Met Asn Lys Lys Asn
100 105 110

Lys Pro Arg Cys Val Cys Ala Pro Asp Cys Ser Asn Ile Thr Trp Lys
115 120 125

Gly Pro Val Cys Gly Leu Asp Gly Lys Thr Tyr Arg Asn Glu Cys Ala
130 135 140

Leu Leu Lys Ala Arg Cys Lys Glu Gln Pro Glu Leu Glu Val Gln Tyr
145 150 155 160

Gln Gly Arg Cys Lys Lys Thr Cys Arg Asp Val Phe Cys Pro Gly Ser
165 170 175

Ser Thr Cys Val Val Asp Gln Thr Asn Asn Ala Tyr Cys Val Thr Cys
180 185 190

Asn Arg Ile Cys Pro Glu Pro Ala Ser Ser Glu Gln Tyr Leu Cys Gly
195 200 205

Asn Asp Gly Val Thr Tyr Ser Ser Ala Cys His Leu Arg Lys Ala Thr
210 215 220

Cys Leu Leu Gly Arg Ser Ile Gly Leu Ala Tyr Glu Gly Lys Cys Ile

225 230 235 240

Lys Ala Lys Ser Cys Glu Asp Ile Gln Cys Thr Gly Gly Lys Lys Cys
245 250 255

Leu Trp Asp Phe Lys Val Gly Arg Gly Arg Cys Ser Leu Cys Asp Glu
260 265 270

Leu Cys Pro Asp Ser Lys Ser Asp Glu Pro Val Cys Ala Ser Asp Asn
275 280 285

Ala Thr Tyr Ala Ser Glu Cys Ala Met Lys Glu Ala Ala Cys Ser Ser
290 295 300

Gly Val Leu Leu Glu Val Lys His Ser Gly Ser Cys Asn
305 310 315

<210> 4
<211> 588
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (377)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (379)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (428)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (439)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (444)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (452)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (461)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (466)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (466)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (468)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (470)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (475)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (485)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (501)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (505)
<223> n equals a, t, g, or c

<400> 4
aattcggcac gagtttcrna gcccsaagcc tctagtctgg gtgtgtacgg agggtctagc 60

ctgggtgtgt acggagggtc tagcctgggt gagtacggag ggtctagcct gggtgagtac 120

ggagggtcta gcctgggtga gtacggagag tctagcctgg gtgtgtatgg aggatctagc 180

ctgggtgagt atggagggtc tagcctgggt gagtatggag ggtctagcct gggtgtgtat 240

ggagggtcta gcctgggrga gtatggaggg tctagcctgg gtgtgtatgg agggtctagc 300

ctgggtgagt atggagggtc tagccttggt gtttacggag ggtctagtct gagtrgtrtt 360

tggggaccts agaacantnt taacctttag ccvagmaanc caggccctta atgaaggcra 420

gaaggttnca ccattcctnc cctnctaaga antcaattty nnaatncntn ttgtnccctt 480

ttggnccttt aagccattta naatncca 508

<210> 5
<211> 466
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (415)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (422)
<223> n equals a, t, g, or c

<220>
<221> misc_feature

<222> (437)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (449)
<223> n equals a, t, g, or c

<400> 5
ggcgacggcg tggagtgcgg ccgggcaag gcgtgccgca tgctgggggg cggccgcgc 60

tgcgagtgcg cgccgactg ctcgggggnkc ccggcgcggt tgcaggtctg cggctnagac 120

ggcgcnacct accgcgacga gtgcgagctg cgcgccgcc gttgccgcgg ccaccccgac 180

ctgagcgtca tgtaccgggg ccgctgccgc aagtcctgtg agcangtggt gtgcccgcgg 240

ccacagtcgt gcgtngtgga ccagacgggc agcgccact gcgtggtgcg tcgaagcggc 300

gccctgccct gtgcvcctcca gccccggcca ggagctttgc ggccaacaac anagttacct 360

aaattttctc gtgccgaatg cgccaaggcc aactgcttcn tgggccggtt ccaannggcg 420

tncgccaagc gggcaanttt cgcaagcanc cctgaaggag ccgcca 466

<210> 6
<211> 337
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (33)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (60)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (92)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (141)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (264)
<223> n equals a, t, g, or c

<400> 6
cttgagtgcg tgtggggacc tcagaacact gtnaccttag cccagcaagc caggcccttn 60

atgaaggcca agaaggctgc caccattccc tnncagtcca agaactccag gttcccaact 120

gcctcttttt gcccctttgc ntcctgtgaa ggccattgag aaatgcccag tgtgcccct 180

gggaaagggc acggcctgtg ctcctgacac gggctgtgct tggccacaga accacccagc 240

gtctcccctg ctgctgtcca cgtnagttca tgaggcaacg tcgcgtggtc ttcagacgtg 300

ggagcagcca gcggcagctc aggaggcagg gaacagt 337

<210> 7
<211> 298
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (5)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (11)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (68)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (76)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (84)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (111)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (186)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (243)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (266)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (272)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (280)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (288)
<223> n equals a, t, g, or c

<400> 7
ggcanagccg nctggtgggt gggatagacc tgcttrcngg acactgagcg cctgatttag 60

ggnccttntn taggaatgcc ccanncccta cccraagacc tattgccggg nagyattcca 120

aanttnnggt ccttgggggn taaancctatt aattattyct antatcaaga gggctgyggc 180

attccntgct ggtaaartcc tgaagaggca tgactgcttc trtaagccnc aagcctctag 240

ttntgggtgt tttarggagg ggtctnagcc tnggggttgtn gtacgggngg ggttctta 298

<210> 8
<211> 186
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (34)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (128)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (155)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (180)
<223> n equals a, t, g, or c

<400> 8
ccggcggagc aaagtccact ctgggggagc tctngcgggg accacgggcc actgctcacc 60

cactggcccc gagggggtg tagacgccaa gactcacgca tgtttgacat ccggagtcct 120

ggagccgngt gtcccagtgg caccactagg tgctnnctgc ctccacagtg gggttcacan 180

ccaggg 186

<210> 9
<211> 308
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (3)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (19)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (24)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (29)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (34)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (38)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (40)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (50)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (83)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (107)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (205)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (220)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (237)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (272)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (342)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (297)
<223> n equals a, t, g, or c

<220>
<221> misc_feature

<222> (308)
<223> n equals a, t, g, or c

<400> 9
ggnagaggtg acaccagcna ggtnctgtnt tggnccantn ctcccacgan ggctcacact 60

cccctccatc tgctttaatg ctncgaatcg cctactgtgg ccctgcntgt aaaccacagc 120

tttcaaacca gctatgggga gaggacaacc cggaggatat tccagcttcc ccggtctggg 180

gtgaaggagt gtggggagct tgggncatcc tcctccagtn tncticagcc cccaggnagt 240

gnctttaanc tgtgggttgc ccagaaaagt gncccttagg tttgttgggt tttaaangga 300

gctttaan 308

<210> 10
<211> 407
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (290)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (298)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (324)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (328)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (386)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (402)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (407)
<223> n equals a, t, g, or c

<400> 10
ggcacgagcc tgggtgtgta cggagggtct agtctgagtg cgtgtggggc ctnagaacac 60

tgtgacctta gcccagcaag ccaggccttc atgaaggcaa gaaggtgccc ccatcccttg 120

ccagccnaag actccagttc cccactgcct ctgtgtgccc ttrgcgtcct gtgaagccat 180

tgagaaatgc ncatgtgccc ctgggaaagg gcacggctgt gtcctgacag ggtgtgtttg 240

cacagcccac caggtttcct gtgtgttagt attatgagga acgtcggtgn ttagggtnga 300

gcagcaggga gttagagcag gatntntccg gggcaagtcc attttgggga tttgcggacn 360

gggcatgtta ccactgcccg aggggntagg gcagttagat tntgaan 407

<210> 11
<211> 139
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (2)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (9)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (31)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (45)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (70)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (105)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (131)
<223> n equals a, t, g, or c

<220>
<221> misc_feature
<222> (139)
<223> n equals a, t, g, or c

<400> 11
angcagggnt nottggtccc ccacaacctt ccccggccag gcctncagac ccagacttca 60

gccagacctn ccttaaccac caatgcagcc ggggcttgcg acaanagcag grgctggttt 120

tggggcagtt nttccangg 139

<210> 12
<211> 25
<212> DNA
<213> Homo sapiens

<400> 12
tcacgccata tgggcttggg gaacc 25

<210> 13
<211> 41
<212> DNA
<213> Homo sapiens

<400> 13
catccgggta cattattacc cgaagttctc ttcctcttct g 41

<210> 14
<211> 40
<212> DNA
<213> Homo sapiens

<400> 14
catcgcggat ccgccatcat gcgtcccggg gcgccagggc 40

<210> 15
<211> 38
<212> DNA
<213> Homo sapiens

<400> 15
catccgggta cctcacacga agttctcttc ctcttctg 38

<210> 16
<211> 40
<212> DNA
<213> Homo sapiens

<400> 16
catcgcggat ccgccaccat gcgtcccggg gcgccagggc 40

<210> 17
<211> 35
<212> DNA
<213> Homo sapiens

<400> 17
tcaccgctcg agcacgaagt tctcttcctc ttctg 35

<210> 18
<211> 40
<212> DNA
<213> Homo sapiens

<400> 18
catcgcggat ccgccaccat gcgtcccggg gcgccagggc 40

<210> 19
<211> 38
<212> DNA
<213> Homo sapiens

<400> 19
catccgggta cctcacacga agttctcttc ctcttctg 38

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/17710

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C07H 21/00; C12N 15/12, 15/62, 15/63

US CL. : 435/69.1, 252.3, 254.11, 320.1, 325; 536/23.1, 23.5, 24.31

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 252.3, 254.11, 320.1, 325; 536/23.1, 23.5, 24.31

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

GENBANK, APS, MEDLINE, EMBASE, BIOSIS, SCISEARCH, CAPLUS
search terms: follistatin

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database GenBank, US National Library of Medicine, (Bathesda, MD, USA), No. N52331, HILLIER et al., 'WashU-Merck EST Project', human, complete record, 30 January 1997.	10, 21
X	Database GenBank, US National Library of Medicine, (Bathesda, MD, USA), No. AA552990, 'National Cancer Institute, Cancer Genome Anatomy Project', human, complete record, 11 August 1997.	10
X	Database GenBank, US National Library of Medicine, (Bathesda, MD, USA), No. N28854, HILLIER et al., 'The WashU-Merck EST Project', human, complete record, 04 January 1996.	10

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"B" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 OCTOBER 1998	06 NOV 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer SCOTT D. PRIEBE Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/17710

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database GenBank, US National Library of Medicine, (Bathesda, MD, USA), No. R74502, HILLIER et al., 'The WashU-Merck EST Project', human, complete record, 05 June 1995.	10
X	Database GenBank, US National Library of Medicine, (Bathesda, MD, USA), No. N75101, HILLIER et al., 'WashU-Merck EST Project', human, complete record, 30 January 1997.	10
E, X	Database GenBank, US National Library of Medicine, (Bathesda, MD, USA), No. AC004156, LAMERDIN et al., 'Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a serine protease cluster', human, partial record -notes and sequence for nucleotides 31801-38820, 19 February 1998.	1-13, 15, 21
A	SHIMASAKI et al. Primary structure of the human follistatin precursor and its genomic organization. Proceedings of the National Academy of Science USA. June 1988, Vol. 85, pages 4218-4222.	1-17, 21, 22
X	Database GenBank, US National Library of Medicine, (Bathesda, MD, USA), No. AA375541, ADAMS et al., 'Initital assessment of gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequencing', human, complete record, 21 April 1997.	21
X	Database GenBank, US National Library of Medicine, (Bathesda, MD, USA), No. AA363365, ADAMS et al., 'Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library', human, complete record, 21 April 1997.	21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/17710

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-17, 21, 22

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/17710

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s)1-17, 21 and 22, drawn to polynucleotides 95% identical to polynucleotide encoding follistatin-3 (SEQ ID NO: 2), cells comprising same on vector, and method of using cell for producing follistatin-3.
Group II, claim(s) 18, 19, 23 and 24, drawn to polypeptides 95% identical to follistatin-3, and method of using same to treat disease.
Group III, claim(s) 20, drawn to antibody that binds to follistatin-3.
Group IV, claim(s) 25, drawn to method of treatment using nucleic acid.
Group V, claim(s) 26, drawn to method of diagnosis by identifying mutation in follistatin-3 DNA.
Group VI, claim(s) 27, drawn to method of diagnosis by measuring amount of follistatin-3 in biological sample.
Group VII, claim(s) 28, drawn to a method of using follistatin-3 protein in ligand binding assay to identify inhibitors/activators of follistatin-3 activity.

The inventions listed as Groups I-IV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The polynucleotides of group I share a certain level of sequence identity to any polynucleotide that would encode follistatin-3 with the amino acid sequence of SEQ ID NO: 2. Most such polynucleotides would not encode a follistatin-3 protein, since most of the changes would be in non-wobble bases and be deletions and insertions that alter the reading frame of the polynucleotide. The polypeptides of group II share a certain level of identity with follistatin-3 with the amino acid sequence of SEQ ID NO: 2, many of these polypeptides would not be follistatin-3, i.e. they would not be active as follistatin-3. Each of groups I and II includes the first named method of using a polynucleotide or polypeptide of groups I and II, which methods require a polynucleotide encoding follistatin-3 or follistatin-3, respectively. Since the majority of polynucleotides and polypeptides of groups I and II, respectively, could not be used for the methods since they would not encode a follistatin-3 or have follistatin-3 activity, respectively, as required in the methods, the products do not share the special technical feature with the methods of using but a small fraction of them. Likewise the methods of groups IV-VII are alternative methods of using a small fraction of the polynucleotides and polypeptides of groups I and II, respectively, that require polynucleotides that encode follistatin-3 or natural mutant variants, or require active follistatin-3. The antibodies of group III are specific for follistatin-3, and have no use for the polynucleotides of group I or most polypeptides encoded by the polynucleotides, and would be specific for only a small fraction of the polypeptides of group II.

Groups I-IV are drawn to multiple products and/or multiple methods of using a product; 37 CFR 1.475(b) does not provide for unity of invention between multiple distinct products or multiple distinct methods of using a product.